OPEN ACCESS

KING'S College LONDON

King's Research Portal

DOI:
10.1016/j.intmar.2019.03.004

Document Version
Peer reviewed version

Link to publication record in King's Research Portal

Citation for published version (APA):
Theodosiou, M., Katsikea, E., Samiee, S., & Makri, K. (2019). A comparison of formative versus reflective approaches for the measurement of electronic service quality. *JOURNAL OF INTERACTIVE MARKETING*, *47*, 53-67. https://doi.org/10.1016/j.intmar.2019.03.004

Citing this paper
Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights
Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

•Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
•You may not further distribute the material or use it for any profit-making activity or commercial gain
•You may freely distribute the URL identifying the publication in the Research Portal

Take down policy
If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 05. May. 2023

A comparison of formative versus reflective approaches for the measurement of electronic service quality

Abstract

Drawing upon the extant literature, we examine the definition, conceptualization, and measurement of electronic service quality (e-SQ) and propose an alternative measurement approach. We posit that customers' perceptions of overall e-SQ are influenced by six proximal antecedents: security/privacy, fulfillment/reliability, website design, customer service, informativeness, and customization. Using three independent samples of real customers of three hotel reservation sites, we assess the proposed measurement approach via appropriate scale development procedures. Findings indicate that a causal structure that considers e-SQ as a distinct construct that is influenced by six proximal antecedents should be preferred over a second-order reflective measurement model. We discuss theoretical implications of our proposed measurement method, make recommendations for service firms, and offer directions for future research.

Keywords: electronic service quality; scale development; formative vs. reflective specification.

Introduction

The Internet has rapidly evolved into an important marketing medium and channel for purveyors of services since its inception. A strong online presence offers enormous opportunities to companies to expand customer base, sales revenues, and profits, contingent on site's effectiveness for attracting qualified visitors, converting them into paying customers, and developing a loyal following (Schlosser, White, and Lloyd 2006). The rapid growth of the Internet in 1990s led to intense competition across platforms which, in turn, center-staged the critical role of service quality as the primary source of competitive advantage in online channels. Superior electronic service quality (e-SQ) enables companies to satisfy customers' needs and preferences, enhance loyalty, and maintain longer term profitable relationships.

Not surprisingly, a growing body of research has developed e-SQ measurement tools and examined its relationship with consumer satisfaction, trust, and loyalty (e.g., Anderson and Srinivasan 2003; Finn, Wang, and Frank 2009; Kim, Jin, and Swinney 2009; Shankar, Smith, and Rangaswamy 2003). There is a broad consensus among researchers that e-SQ enhances customer loyalty (Chang and Chen 2008; Harris and Goode 2004). However, researchers also acknowledge that the nature and the strength of the relationships between e-SQ, e-loyalty, and their correlates are context-specific and vary across different cultures, product categories, and websites (Gupta and Kabadayi 2010). Therefore, further research is essential to uncovering a comprehensive set of e-loyalty antecedents within different empirical settings. Another research stream focuses on the conceptual definition and measurement of the key dimensions of e-SQ and the development of psychometrically sound scales (e.g., Bauer, Falk, and Hammerschmidt 2006; Fassnacht and Koese 2006; Parasuraman, Zeithaml, and Malhotra 2005; Wolfinbarger and Gilly 2003). These efforts have led to different conceptualizations of e-SQ.

Nevertheless, a number of critical issues heighten the importance of further research on e-

SQ. Our goal in this study is to highlight and address three important research gaps in the extant literature. *First*, some researchers treat individual dimensions of e-SQ as direct drivers of satisfaction, trust, or loyalty (e.g. Bart, Shankar, Sultan and Urban 2005; Srinivasan, Anderson, and Ponnavolu 2002). However, customers' attitudes and behavior are most likely influenced by overall service quality perceptions (Brady and Cronin 2001). Therefore, there is a need to develop a robust measurement approach for overall e-SQ. Our objective in this study is to demonstrate that customers' perceptions of overall e-SQ is a distinct construct which can be measured directly and independently of its causes. *Second*, most studies conceptualize e-SQ as comprising of reflective indicators (Ladhari 2010). Some have argued that an e-SQ scale comprised of reflective indicators leads to model misspecification and ultimately to biased results (Collier and Bienstock 2006). To the best of our knowledge, no previous study has explicitly compared a measurement approach based on the use of reflective indicators with the alternative which is based on the use of formative indicators. This study assesses the appropriateness of each measurement approach by examining the relevant theory and employs robust structural equation modeling procedures to empirically compare a reflective vs. a formative specification for e-SQ. Based on our findings, we propose a measurement approach that rests on defining and measuring overall e-SQ as a distinct entity, and treat individual service quality dimensions as proximal antecedents of this construct. *Third*, most previous research has focused on online shopping for goods, largely ignoring electronic services (Fassnacht and Koese 2006).[1] Thus, a universal measurement model applicable across different kinds of electronic services is yet to emerge (Fassnacht and Koese 2006; Ladhari 2010). Currently, there is no consensus either in the number and nature of e-SQ dimensions or the exact structural relationship

[1]We distinguish between online shopping and electronic services in that the former typically deals with purchases of tangible goods, whereas the latter group generally deals with arrangements involving intangible goods, such as hotel or airline reservations.

between these dimensions and the overall e-SQ construct. This study also contributes to the literature by using data from three independent samples drawn from the real customer data bases of three different hotel reservation sites.

We accomplish our objectives in three separate studies. Studies 1 and 2 involve the development and empirical validation of a distinct measure for e-SQ as well as six proximal antecedents of customers' perceptions of overall e-SQ. Study 3, in turn, further assesses the external validity of the scales developed in Studies 1 and 2, using data gathered from customers of a different service-based website. In the following sections, we first discuss the research background and hypotheses. We then present Studies 1and 2, which involve an iterative procedure for the conceptual development and empirical assessment of e-SQ and its proximal antecedents. Next, we present Study 3, which replicates the measurement scales developed in Studies 1 and 2 using a new sample of customers, and provides an empirical test for an expanded model of e-SQ. Using data for Study 3, we also perform a direct comparison of a reflective vs. a formative measurement model specification for e-SQ. We conclude with theoretical and managerial implications, limitations, and future research.

Theoretical Background

Three issues regarding e-SQ are key concerns in this investigation. These include defining e-SQ, identifying its underlying dimensions, and employing a robust approach for its measurement (Zeithaml, Parasuraman, and Malhotra 2002). Regarding a definition, Zeithaml, Parasuraman, and Malhotra (2002, p. 363) defined e-SQ as "the extent to which a website facilitates efficient and effective shopping, purchasing, and delivery of products and services." A more generic definition is provided by Fassnacht and Koese (2006, p. 25) who assert that "the quality of electronic services is the degree to which an electronic service is able to effectively and

efficiently fulfill relevant customer needs." Accordingly, within the context of this research, we define e-SQ as the perceptions of existing users regarding the overall quality of hotel reservation websites.

A review of the literature reveals a consensus among researchers regarding the multidimensionality of e-SQ, with multiple studies focused on identifying its relevant dimensions. Thus, a range of e-SQ models using different quality dimensions have been proposed and several reviews of e-SQ scales have been published.[2] Given the volume of published works addressing e-SQ, it would be redundant to present yet another comprehensive review of this literature as a precursor to our investigation (see Appendix 1 for a summary of key e-SQ studies). However, four important limitations of the extant literature motivate us to reassess the aforementioned critical issues concerning e-SQ construct (i.e., definition, conceptualization, and measurement) and serve as the basis for the need for further development of this scale and, hence, this investigation.

First, the majority of early e-SQ studies examine website interface, and ignore the complete purchase experience (e.g., Collier and Bienstock 2006). Second, the terminology used in different studies is seemingly convoluted (Holloway and Beatty 2008). In fact, most lack a clear definition for e-SQ construct, even though the same terms have been used to refer to different sub-dimensions of e-SQ across studies (Fassnacht and Koese 2006). Third, rigorous tests of the psychometric properties of their newly-developed scales have been scarce (Ladhari 2010). Finally, e-SQ studies have often relied on convenient samples of students. Students'

[2]These include SITEQUAL (Yoo and Donthu 2001), PIRQUAL (Francis and White 2002), WEBQUAL (Loiacono, Watson and Goodhue 2002), eTailQ (Wolfinbarger and Gilly 2003), E-S-QUAL (Parasuraman, Zeithaml, and Malhotra 2005), eTransQual (Bauer, Falk, and Hammerschmidt 2006), and ER-SERVCOMPSQUAL (Rossiter 2009). Representative reviews and assessments of the e-SQ literature include Fassnacht and Koese (2006), Holloway and Beatty (2008), Ladhari (2010), Lin and Hsieh (2011), Rowley (2006), Tan, Benbasat, and Cenfetelli (2013), and Parasuraman, Zeithaml, and Malhotra (2005).

evaluations are usually limited to the characteristics of a particular website and do not consider the whole purchase process (Parasuraman, Zeithaml, and Malhotra 2005). Moreover, the use of student populations limits the generalizability of the scales and reduces their applicability to broader populations of online users (Ladhari 2010). Overall, the extant literature leads to the conclusion that available e-SQ scales are context-specific, suffer from poor external validity, and cannot be used for different online services (Fassnacht and Koese 2006; Ladhari, 2010). These conditions serve as the basis for a new e-SQ scale that follows rigorous psychometric scale construction procedures for use in a given situation.

e-SQ Measurement Specification

In general, two measurement perspectives regarding the structural relationship between the latent construct and its indicators exist: one based on the use of reflective indicators, and another based on the use of formative indicators (Diamantopoulos and Winklhofer 2001; Jarvis, MacKenzie, and Podsakoff 2003). Four criteria determine whether a construct is formative or reflective: *(a)* the direction of causality, *(b)* the interchangeability of the indicators, *(c)* the extent of covariation among the indicators; and *(d)* whether measures have the same antecedents and consequences (Jarvis, MacKenzie, and Podsakoff, p. 203). In a reflective-indicator model, the direction of causality is from the construct to the indicators, and changes in the underlying construct are expected to cause changes in the indicators. Moreover, indicators are interchangeable, share a common theme, and are expected to covary; therefore, dropping an indicator should not alter the conceptual domain of the construct. On the other hand, in a formative-indicator model, the direction of causation is from the indicators to the construct, and changes in the indicators are expected to cause changes in the underlying construct. Indicators need not be interchangeable or share a common theme, do not need to covary, and dropping an indicator may alter the

conceptual domain of the construct. Finally, in reflective models, indicators are required to have the same antecedents and consequences, something which is not necessary in formative models (Jarvis, MacKenzie, and Podsakoff 2003, p. 203).

Most studies conceptualize e-SQ as a higher-order reflective construct (e.g., Bauer, Falk, and Hammerschmidt 2006), whereas other studies adopt a formative model (e.g., Collier and Bienstock 2006; Rossiter 2007, 2009). Surprisingly, most studies do not provide a justification for choosing a reflective or formative scale. To arrive at a more balanced, thorough, and seasoned decision about the measurement specification of e-SQ, we next provide a theory-driven solution to address this issue, followed by an investigation of empirical complexities of formative-indicator models.

Theoretical Considerations

Most studies that address the formative vs. reflective issue treat e-SQ as a formative-indicator construct (e.g., Collier and Bienstock 2006). *First*, it is reasonable to expect each e-SQ dimension to cause overall e-SQ and not vice versa. For example, a user's assessment of a website's quality in terms of *customer service*, *privacy*, or *design* contributes to forming perceptions about the overall e-SQ of the site. *Second*, e-SQ dimensions are not interchangeable. For example, *security/privacy* captures a unique aspect of the conceptual domain of the construct which cannot be substituted by other dimensions like *site design* or *customization*. *Third*, the different e-SQ dimensions are not required to covary. *Finally*, there is no evidence that the different e-SQ dimensions have the same antecedents and consequences. These theoretical criteria lead to the conclusion that e-SQ should be specified as a (higher-order) formative-indicator construct.

Empirical Considerations

Multiple studies assert that many constructs are incorrectly operationalized as reflective, whereas a formative scale would have better captured the domain. A likely explanation for this trend is the fact that a reflective specification is more convenient (e.g., Rossiter 2002). In contrast, the estimation of formative models is challenging. A major limitation of formative-indicator models is that they are under-identified and can be estimated only by following specific approaches. Jarvis, MacKenzie, and Podsakoff (2003, p. 213) propose that to achieve measurement model identification, each construct with formative indicators must emit paths to: *(1)* at least two unrelated latent constructs with reflective indicators; *(2)* **at least two theoretically appropriate reflective indicators**; or *(3*) one reflective indicator and one latent construct with reflective indicators. However, the two options that involve additional latent constructs with reflective indicators (i.e., options 1 and 3) have serious limitations, which relate to the error term or "surplus meaning" associated with formative constructs. In light of these problems, Jarvis, MacKenzie, and Podsakoff's (2003) recommendation for solving the identification problem by adding at least two reflective indicators to a formatively specified construct is likely the best approach. Three advantages are evident in this approach: "*(a)* the formative construct is identified on its own and can go anywhere in the model; *(b)* one can include it in a confirmatory factor model and evaluate its discriminant validity and measurement properties; and *(c)* the measurement parameters should be more stable and less sensitive to changes in the structural relationships emanating from the formative construct" (Jarvis, MacKenzie, and Podsakoff 2003, p. 213). The two reflective indicators should provide a global assessment of the construct (e.g., overall e-SQ, overall job satisfaction, etc.).

Following this logic, Wilcox, Howell, and Breivik (2008, pp. 1226-27) argue that "if the

researcher is able to obtain two reflective indicators, obtaining at least one more reflective indicator (but preferably several more) would result in a testable reflectively measured construct, and the problem of developing formative measures vanishes." Accordingly, we propose that customers' perceptions of overall e-SQ consist of a distinct construct which can be measured directly. Studies that define overall e-SQ as a distinct construct (measured directly and independently of its underlying dimensions), provide support for the veracity of this position (e.g., Barrutia and Gilsanz 2013; Fassnacht and Koese 2006; Montoya-Weiss, Voss, and Grewal 2003). Another critical issue is whether individual quality dimensions (e.g., security/privacy, fulfillment/reliability, etc.) can or should be considered as causes of overall e-SQ or integral aspects of this construct. Wilcox et al. (2008) and Edwards (2011) argue that formative indicators cannot be both causes and integral parts of a construct, since causes and effects should be distinct entities. Therefore, it is plausible to consider e-SQ dimensions as causes of customers' perceptions of overall e-SQ, vis-à-vis defining them as formative indicators of e-SQ. However, the dimensions of a construct cannot be treated as causes of the construct in the same way as any other causes (MacKenzie, Podsakoff, and Podsakoff 2011). Rossiter (2002) addresses this problem by making a distinction between proximal antecedents and the more remote causes, which is consistent with the goals of our research. As such, we conceptualize e-SQ dimensions as proximal antecedents of overall e-SQ.

The Identification of e-SQ Proximal Antecedents

First, we identify the e-SQ dimensions whose conceptual relevance and empirical significance have been established. There is no consensus on the relevant e-SQ dimensions, but the literature reveals that recent studies have relied on particular e-SQ dimensions. These include fulfillment/reliability, privacy/security, website design, and customer service (e.g., Holloway and

Beatty 2008; Parasuraman, Zeithaml, and Malhotra 2005). Another important e-SQ dimension relates to the provision of in-depth, accurate, relevant, and timely information (e.g., Collier and Bienstock 2006; Fassnacht and Koese 2006; Wolfinbarger and Gilly 2003). There is also ample support that informativeness is also a component of e-SQ (e.g., Montoya-Weiss, Voss, and Grewal 2003; Wolfinbarger and Gilly 2003). Finally, website customization is a quality dimension that is also frequently discussed (Rowley 2006), but is absent from the vast majority of available e-SQ scales. This is surprising, considering that the opportunity for personalization is a key feature of online channels and plays a crucial role in online customer relationship management programs. Moreover, empirical evidence indicates that customization is a significant correlate of online customer loyalty (Srinivasan, Anderson, and Ponnavolu 2002) and satisfaction (Chang and Chen 2008). These insights support the inclusion of customization as a sixth dimension in our model of the proximal antecedents of e-SQ.[3] Below, we briefly describe the six dimensions of quality as proximal antecedents of e-SQ scale in our investigation.

Privacy and security are major evaluative criteria in online services (Holloway and Beatty 2008; Parasuraman, Zeithaml, and Malhotra 2005). Privacy includes the protection of personal information, protecting anonymity, and providing informed consent. Security involves protecting users from the risk of fraud and financial loss from the misuse of their credit card or other financial information. ***Fulfillment and reliability*** capture issues relating to on-time and accurate delivery of the product or service, accurate product/service presentation, and error-free transactions (Zeithaml, Parasuraman, and Malhotra 2002). ***Informativeness*** refers to the presentation of accurate, timely, and complete information on a website and includes the availability of research facilities (Collier and Bienstock 2006; Wolfinbarger and Gilly 2003).

[3]Given that we aim to develop context-free scales of e-SQ and its antecedents which are applicable to different kinds of electronic services, quality-related logistics aspects of shopping (e.g., shipment tracking, delivery condition, timeliness of delivery) were not considered.

Customization is the ability of an e-service provider to tailor and adapt products, services, and the transactional environment to individual customers (Srinivasan, Anderson, and Ponnavolu 2002). People value personalized experiences (Feldman and Lynch 1988), and customization increases the probability that customers will find something that they wish to buy. ***Website design*** includes the elements of a customer's experience with the website such as navigation, site organization and layout, graphic design, and visual appearance (Collier and Bienstock 2006; Holloway and Beatty 2008). Finally, ***customer service*** is defined as receptive, helpful, and willing service that rapidly responds to customer inquiries (Wolfinbarger and Gilly 2003).

Hypotheses Development

As previously discussed, we propose that customers' perceptions of e-SQ is a distinct construct which can be defined independently of its causes; the various e-SQ dimensions are considered as proximal antecedents of this construct. In line with this conceptualization, we assert:

> *H1*: (a) Security/privacy, (b) fulfilment/reliability, (c) informativeness, (d) customization, (e) website design, and (f) customer service have positive impacts on customers' perceptions of overall e-SQ.

To provide a comprehensive assessment of the psychometric properties of the proposed e-SQ scale, a test of nomological validity is required (Lewis, Templeton, and Byrd 2005). To assess nomological validity, we introduced three additional constructs in Study 3, namely trusting beliefs, satisfaction, and loyalty intentions. Trusting beliefs and customer satisfaction represent important sought-after outcomes of e-SQ and therefore are appropriate for testing its nomological validity. Although not a central goal, we also assess the impact of trusting beliefs and customer satisfaction on customer loyalty intentions. Our conceptual model is tested in Study 3 (Figure 1) and the following hypotheses summarize the rationale for each relationship.

…Insert Figure 1 about here…

e-SQ and trusting beliefs. The critical role of trust in successful business relationships has long been established (Doney and Cannon 1997; Morgan and Hunt 1994). Trusts consists of two interrelated components: trusting beliefs, i.e., expectations about an exchange partner's trustworthiness (Moorman, Deshpandé, and Zaltman 1993); and trusting behavior, i.e., willingness to depend on a vendor (Mayer, Davis, and Schoorman 1995). Trust is of paramount importance in online markets (Harris and Goode 2004), and given that customers interact with the firm online, a website that conveys a high level of service is the best way to build and enhance trusting beliefs. Research has demonstrated a positive relationship between e-SQ and customer trust (e.g., Chen and Dibb 2010; Ha and Stoel 2009). Therefore, we posit:

H2: e-SQ has a positive influence on customers' trusting beliefs.

e-SQ and satisfaction. e-SQ has a positive influence on consumers' online satisfaction (e.g., Evanschitzky et al. 2004). Specifically, satisfaction is the pleasurable fulfillment of a service (Oliver 1999) which is more likely to occur when the experience of an online customer exceeds his/her expectations (Oliver 1997). Thus, higher quality levels of an e-service lead to higher satisfaction (Carlson and O'Cass 2010; Chang and Chen 2008). We thus propose:

H3: e-SQ has a positive influence on customers' satisfaction.

Trusting beliefs and loyalty intentions. Loyalty intentions reflect a customer's keenness to recommend a website to others, as well as his/her intentions to stay with it (Parasuraman, Zeithaml, and Malhotra 2005). Customers must trust a website to remain loyal online because of the absence of personal contact and concerns about security and fraud (Harris and Goode 2004). Trust is a key determinant of loyalty in online environments (e.g., Gupta and Kabadayi 2010; Harris and Goode 2004; Kim, Jin, and Swinney 2009). Thus, we assert:

H4: Consumers' trusting beliefs has a positive influence on their loyalty intentions.

Satisfaction and loyalty intentions. Customer satisfaction and loyalty are inherently linked

(Oliver 1999). Research has demonstrated a positive relationship between e-satisfaction and e-loyalty, albeit generally within tangible product contexts (Anderson and Srinivasan 2003; Kim, Jin, and Swinney 2009). The literature offers some evidence of this relationship in other contexts, but its inclusion in our model is necessary and affords us the opportunity to re-examine it in an online, uniform, service-based context. Therefore, we propose:

H_5: Customer satisfaction has a positive influence on loyalty intentions.

Study 1: Initial Assessment of the e-SQ Measurement

We aim to develop appropriate measurement scales for overall e-SQ and its proximal antecedents in Study 1. Importantly, we perform an initial exploratory assessment of the psychometric properties of the respective measurement scales.

Sample and Data Collection Procedure

We used an online hotel reservation site for data collection purposes. The focal website operates as a hotel meta-search site and provides customers with the best deals online. As is commonplace with many such sites, the reservation is completed on hotels' own websites. Our sample consisted of customers who made a specific hotel reservation during the 12-month period preceding this study (the standard in the tourism/hospitality industry and used by researchers). This time period is a full fiscal year within which an average customer is expected to have made at least one online hotel booking. The focal site provided all customer-related and contact information.

We drew a random sample of 1,000 customers from the available pool for data collection. We developed a structured questionnaire consisting mostly of close-ended questions, and employed a professional service to develop an online version of the questionnaire. An e-mail

with a cover letter explaining the purpose of the study and a URL link to the questionnaire was sent to all respondents. We received 139 usable responses for a response rate of 13.9%.

Measures

Our constructs were developed based on the extant literature, and measurement items were adapted to our study context. We measured security/privacy, fulfillment/reliability, customer service, informativeness, and satisfaction using scales adapted from Wolfinbarger and Gilly (2003). For customization, we adapted items from Wolfinbarger and Gilly (2003) and Srinivasan, Anderson, and Ponnavolu (2002). Our website design measure is based on the scale developed by Bart et al. (2005), and adopts a broader definition of the term "navigation" to cover important website design aspects like appearance, layout, and ease of use. Website design also includes two items referring to the navigation structure obtained from Montoya-Weiss, Voss, and Grewal (2003). Finally, overall e-SQ was consisted of a four-item scale (Montoya-Weiss, Voss, and Grewal 2003).

We conducted personal interviews with focal hotel reservation website customers to ensure a robust operationalization of constructs. We administered the questionnaire to several researchers familiar with online marketing, e-SQ, and tourism to assess its face validity. The final version of the questionnaire was pre-tested using a random sample of 100 customers, and no problems were observed with respect to measures, their clarity, or the questionnaire length.

Analysis and Results

Study 1 data were only subjected to exploratory analysis to purify our initial measures. We used a combination of exploratory factor analysis (EFA) and item-to-total correlations. We used principal components with varimax rotation. Following recommended standards, we dropped

items with factor loadings below .50 on all factors, as well as those with loadings greater than .50 on two or more factors, or item-to-total correlations less than .40. Retained items loaded heavily on their *a priori* specified factor and had low loadings with the other factors (Appendix 2).

Study 2: Measurement Validation

In Study 2 we assess the validity and reliability of the measurement scales developed in Study 1, using confirmatory factor analysis (CFA). To insure robust results, we rely on a *second* hotel reservation site for data in Study 2.

Sample and Data Collection Procedure

We followed identical sampling and data collection procedures as in Study 1. Data were obtained via an online questionnaire using a different hotel reservation site. Our sampling frame consisted of all customers who made a hotel reservation during the 12-month period preceding Study 2. The site provided customer contact information, and e-mails with a URL link to the site hosting the questionnaire were sent to a random sample of 1,000 customers. To encourage participation, all those who completed the questionnaire were automatically entered in a drawing for a two-night free accommodation for two at a luxury hotel. We received 262 usable questionnaires, for a response rate of 26.2%, equally split between males and females (n = 131).

Analysis and Results

Overall e-SQ and the six proximal antecedents of this construct were measured using the same items employed in Study 1 and retained after the scale purification procedures (see Table 1). We performed CFA using the EQS statistical package (Bentler 1995) to assess constructs'

convergent and discriminant validity. We used the elliptical reweighted least square procedure (ERLS) which provides unbiased parameter estimates for multivariate normal and nonnormal data (Sharma, Durvasula, and Dillon 1989). In estimating the CFA model, each item was restricted to load on its *a priori* specified factor, with the underlying factors being permitted to correlate (Gerbing and Anderson 1988). To assess the goodness of fit of the estimated model, in addition to the chi-square statistic we used three common fit indices (comparative fit index: CFI, the non-normed fit index: NNFI, and the root mean square error of approximation: RMSEA) (Hu and Bentler 1999; McDonald and Ho 2002) which are also least sensitive to sample size variations (Fan, Thompson, and Wang 1999).

Our chi-square is significant ($\chi^2_{506} = 1114.71$, $p < .01$), but alternative fit indices suggest adequate model fit: CFI = .98; NNFI = .98; RMSEA = .068) (Sharma et al. 2005). Moreover, all loadings were large and significant, indicating convergent validity (Table 1). We assessed discriminant validity using the most restrictive test provided by Fornell and Larcker (1981); for all possible pairs of constructs in our study, the shared variance was lower than the average variance extracted for individual constructs. These results provide evidence of discriminant validity. To assess construct reliability, we estimated Cronbach's alpha coefficient, composite reliability score, and average variance extracted. For all constructs, the values obtained for these indices exceed the standards recommended in the literature (Table 1), thus demonstrating the reliability of our measurement scales.

…Insert Table 1 About Here…

Study 3: Measurement Replication and Nomological Validation

Our key goal in this investigation is to confirm the proposed e-SQ measurement approach as comprising of a distinct overall measure of e-SQ and six proximal antecedents. In addition, we

assess the nomological validity of the e-SQ scale by examining its impact on customers' trusting beliefs (H2) and satisfaction (H3). We also investigate the influence of trusting beliefs (H4) and customer satisfaction (H5) on customers' loyalty. Further, we compare our proposed formative specification of e-SQ with an alternative measurement and structural model, in which e-SQ is specified as a second-order reflective construct. Finally, we compare a full mediation structural model, in which the proximal antecedents of e-SQ influence customer trusting beliefs and satisfaction only indirectly (through the mediating effect of e-SQ), with a partial-mediation model, in which the six proximal antecedents of e-SQ can influence trusting beliefs and satisfaction both directly and indirectly (Figure 1).

Sample and Data Collection Procedure

We used a *third* online hotel reservation site for Study 3. As in Studies 1 and 2, our sampling frame consisted of all customers who made a hotel reservation during the 12-month period preceding Study 3. We e-mailed 1,000 random customers, included a URL link, stating the objectives of the study and the importance of their contribution to its success. We received 228 usable questionnaires, a response rate of 22.8% (male = 58.3%, college graduates = 71.4%).

Measures

We used the validated (Study 2) measurement scales for overall e-SQ and its six proximal antecedents. Customer trusting beliefs, satisfaction, and loyalty intentions were three additional constructs to allow an assessment of the nomological validity of our e-SQ scale. We obtained the scale for trusting beliefs from Schlosser, White, and Lloyd (2006). For customer satisfaction, we used a three-item scale obtained from Wolfinbarger and Gilly (2003). Our measure for loyalty intentions was drawn from Parasuraman, Zeithaml and Malhotra (2005).

Measurement Model Estimation

Studies 1 and 2 provide a strong support for validity and reliability of our proposed measurement model for e-SQ. Concurrently, the large number of items used in measuring e-SQ and its proximal antecedents (34 in total) is admittedly a key constraint that may obscure the use of the specific scale in future research.[4] Therefore, we sought to further refine our measurement model by removing certain items which may perform well in absolute terms (i.e., satisfy the recommended standards) but are less effective indicators of their underlying construct. In this respect, all items were subjected to EFA, and stricter criteria were applied to assess their significance. Thus, items with cross loadings of 0.4-0.5 or loadings substantially lower than those of the other items measuring the same construct were considered for elimination. Using this procedure, in combination with item-to-total correlation analysis, we dropped nine items (Table 1). Subsequent analyses were conducted with the remaining items.

Following Anderson and Gerbing (1988), we first estimated the measurement model. The CFA model included our six proximal antecedents, e-SQ, trusting beliefs, satisfaction, and loyalty intentions. All measurement items were restricted to load on their *a priori* specified factor, with the underlying factors being permitted to correlate. Trusting beliefs was specified as a second-order factor comprising three dimensions: ability, benevolence, and integrity. The chi-square value of 1866.77 with 941 degrees of freedom is significant ($p < .01$), indicating the absence of an exact fit. However, all other fit indices suggest adequate model fit: CFI = .98; NNFI = .97; RMSEA = .066. All factor loadings (first- and second-order) are large and significant indicating convergent validity (Table 1). Moreover, reliability estimates satisfy the

[4]We thank an anonymous reviewer for raising this concern.

recommended standards. Furthermore, for all possible pairs of constructs the shared variance is lower than the average variance extracted for individual constructs, providing evidence of discriminant validity. Overall, these findings demonstrate adequate measurement properties for Study 3 scales (Table 2).

…Insert Table 2 About Here…

Structural Model Estimation

We tested our hypotheses by estimating the structural model shown in Figure 1. Table 3 presents parameter estimates, *t*-values, and fit statistics. The hypothesized model has a good fit to the data ($\chi^2_{(961)}$ = 2013.91, $p < .01$; CFI= .97, NNFI= .97, and RMSEA= .069). Importantly, standardized coefficients and corresponding *t*-values (Table 3) provide support for all but one of our hypotheses. Specifically, we find that security/privacy (*H1a*: β = .13, *t* = 2.34), fulfillment/reliability (*H1b*: β = .37, *t* = 6.27), informativeness (*H1d*: β = .71, *t* = 11.18), website design (*H1e*: β = .11, *t* = 1.93), and customization (*H1f*: β = .15, *t* = 2.82) have a significant positive influence on the formation of customers' overall perceptions of e-SQ. On the other hand, customer service is not a significant antecedent of e-SQ (*H1c*: β = .04, *t* = .71). Furthermore, customers' perceptions of e-SQ have a significant positive effect on customer trusting belief (*H2*: β = .52, *t* = 4.97) and customer satisfaction (*H3*: β = .65, *t* = 7.18). In turn, customer trusting beliefs (*H4*: β = .27, *t* = 4.09) and customer satisfaction (*H5*: β = .74, *t* = 8.79) have a positive impact on customer loyalty intentions. The R^2 values obtained for the dependent variables were as follows: e-SQ 84%; trusting beliefs 42%; satisfaction 58%; and loyalty intentions 82%.

…Insert Table 3 About Here…

Estimation of Partial Mediation Models

The structural model (Figure 1), represented by solid lines, is a full mediation model that

captures only the indirect effects of the six proximal antecedents of e-SQ on trusting beliefs and satisfaction, through the mediating effect of e-SQ. To investigate whether these antecedents of e-SQ also directly affect trusting beliefs and satisfaction, we compared the full mediation model with two partial mediation models. The dashed lines in Figure 1 represent the direct influence of the antecedents of e-SQ on trusting beliefs and satisfaction. Because of the large number of parameters included in our structural model (and sample size restrictions), it was not possible to simultaneously test the direct effects on both trusting beliefs and satisfaction. Therefore, two separate partial mediation models were estimated. For trusting beliefs, results indicate that the partial mediation model has a significantly better fit ($\Delta\chi^2 = 103.45$, Δd.f. = 6, $p<.001$). Moreover, two of the parameters capturing the direct effects are statistically significant. Specifically, customer service ($\beta = .19$, $t = 2.47$) and customization ($\beta = .46$, $t = 4.70$) emerge as significant antecedents of trusting beliefs. Although customer service has no significant influence on e-SQ (H1c was not supported), it has a direct effect on customers' trusting beliefs. Moreover, customization is also a critical factor since it influences trusting beliefs both directly and indirectly through the mediating effect of e-SQ. The second partial mediation model captured the direct influences of the six proximal antecedents of e-SQ on customer satisfaction. This model does not have a significantly better fit than the full mediation model: $\Delta\chi^2 = 12.19$, Δd.f. = 6, $p>.05$. Furthermore, none of the antecedents of e-SQ has a significant direct influence on customer satisfaction. Given these results, the addition of twelve new parameters in the structural model for capturing the direct effects of the antecedents of e-SQ on trusting beliefs and satisfaction is not warranted. Therefore, we recommend the full mediation model.

Assessing a Reflective Specification of e-SQ

The formative vs. reflective scales literature asserts that the choice (reflective versus formative

specification) should be theoretically-driven. Previous studies have developed key criteria which can assist researchers in determining whether a construct should be specified as formative or reflective (e.g., Diamantopoulos and Winklhofer 2001). Following a thorough review and assessment of these guidelines, we conclude that e-SQ should be specified as a higher-order formative construct. However, as previously noted, we hold that: (1) e-SQ is a distinct construct which can be measured directly and (2) various dimensions of e-SQ should be considered as proximal antecedents (instead of formative indicators). Nonetheless, to compare the two, the estimation of an alternative measurement and structural model that treats e-SQ as a second-order reflective construct has merit.[5]

In doing so, we first developed a reduced version of our measurement model that included only seven constructs: the six proximal antecedents of e-SQ and the overall e-SQ that was measured with three reflective indicators. Using these constructs, we compared: (1) a formative model in which overall e-SQ is specified as a second-order formative construct that is caused by six first-order factors that serve as its formative indicators; and (2) a reflective model in which e-SQ is specified as a second-order reflective construct comprised of six first-order dimensions. The reflective model included also the overall e-SQ construct, which was allowed to correlate freely with the second-order e-SQ construct. Thus, the formative and the reflective measurement models included exactly the same observed variables and employed the same covariance matrix. The results obtained from the estimation of the two models reveal that the formative model has a better fit to the data (Formative model: $\chi^2_{(254)}$=561.80; CFI=97; NNFI=.98; RMSEA=.073 vs. Reflective model: $\chi^2_{(268)}$=693.35; CFI=96; NNFI=.96; RMSEA=.084). Moreover, in the reflective model, the loadings of the six first-order dimensions

[5]We thank the editor and the reviewers for this suggestion.

on e-SQ are large and statistically significant. In the formative model, the strength and significance of the path coefficients associated with the six proximal antecedents (formative indicators) of overall e-SQ are identical to those obtained from the estimation of the structural model (Table 3), and indicate that with the exception of *customer service*, all other dimensions have a significant influence on overall e-SQ. Overall, the empirical comparison of a formative vs. a reflective measurement model, along with the theoretical considerations previously discussed, lead to the conclusion that e-SQ should be specified as a second-order formative construct caused by six first-order formative indicators (considered in this study as "proximal antecedents").

Table 4 presents the results from the estimation of the structural model, with e-SQ specified as a second-order reflective construct. The findings indicate that e-SQ has a strong influence on customer trusting beliefs ($\beta = .76$, $t = 5.98$) and satisfaction ($\beta = .76$, $t = 7.30$), providing evidence for the construct's nomological validity. The fit of the model is slightly worse compared to the respective formative model ($\chi^2_{(847)} = 1923.70$, $p < .01$; CFI = .96, NNFI = .96, and RMSEA = .075). Notably, however, the parameter estimates are larger than those obtained from the estimation of the formative model. Nonetheless, before concluding that the reflective model is superior in this respect, we need to consider that the two models are not directly comparable. In the formative model, the six antecedents are causing e-SQ, whereas in the reflective model, the e-SQ construct is causing the six dimensions. This is a fundamental difference between the two models. It should also be taken into consideration that, in their influential study, Jarvis et al. (2003, p. 212) concluded that "paths emanating from a construct with a misspecified measurement model are likely to be substantially inflated, thus leading to Type I error." Therefore, the large parameters estimates obtained from the estimation of the reflective model may be due to this kind of inflation.

…Insert Table 4 About Here…

Overall, our empirical results indicate that the specification of e-SQ as a second-order reflective construct meets the relevant validity and reliability standards (although the fit of the respective measurement and structural models is slightly worse compared to the respective formative models). However, as previously discussed, a thorough assessment of the structural characteristics of the e-SQ construct against a number of fundamental theoretical considerations leads to the definite conclusion that e-SQ should be modelled as formative.

Discussion and Implications

Our main objective in this study was to develop and empirically validate a psychometrically robust measurement model of e-SQ and its proximal antecedents. Leveraging the literature on formative vs. reflective scales, we proposed that customers' perceptions of overall e-SQ can be conceptualized as a distinct construct which can be measured directly and independent of its causes. Moreover, we posited that dimensions of e-SQ can be conceptualized as proximal antecedents of e-SQ. We also proposed the six literature-driven antecedents of e-SQ: security/privacy, fulfillment/reliability, informativeness, customization, website design, and customer service. Our findings indicate that with the exception of customer service, all other quality aspects are indeed significant drivers of customers' perceptions of overall e-SQ. A closer analysis of the results reveals that informativeness emerges as the strongest antecedent of e-SQ, which highlights the importance of providing customers with timely, accurate, and detailed information. Although e-SQ measurement research is still developing, certain quality aspects consistently emerge as critical determinants of customers' e-SQ assessment. Specifically, the six proximal antecedents of e-SQ in this study include factors that have been previously identified as

the strongest drivers of customer evaluations of the quality for online service providers (e.g., Ha and Stoel 2009; Holloway and Beatty 2008; Lin and Hsieh 2011). In sum, this aspect of our study integrates well with existing research and helps advance the development of the body of knowledge pertaining with the conceptualization and measurement of e-SQ.

Furthermore, our results identify the proposed measurement approach and resulting causal structure for e-SQ as superior to the alternative that treats this construct as a second-order reflective scale. First, our theoretical arguments support the specification of e-SQ as a formative construct. Second, our empirical findings demonstrate that a comparison of a structural model that treats e-SQ as a distinct construct with six proximal antecedents has a slightly better fit than a similar model treating e-SQ as a second-order reflective construct comprised of six first-order dimensions, further corroborating our theoretical argumentation. Importantly, the six proximal antecedents explain 84% of the variation of the overall e-SQ construct, which implies that each antecedent contributes significantly in forming consumers' perceptions regarding the overall quality of an online service provider. Notably, formatively-specified models are based on a multiple regression (Diamantopoulos and Winklhofer 2001) and therefore multicollinearity among the (formative) indicators poses the same problems as in that statistical technique (Mackenzie, Podsakoff, and Jarvis 2005). If the six proximal antecedents/dimensions of e-SQ were interchangeable (a key criterion for identifying a construct as reflective), only a few of them would contribute towards explaining the variance of the construct, leaving the coefficients of the other dimensions to be insignificant (Law and Wong 1999). However, in our study only one antecedent of e-SQ (customer service) is insignificant. Considering these results, along with the relatively high R^2 and the fact that the formative model has a better fit than the reflective model, it would be reasonable to conclude that the six proximal antecedents/dimensions of e-SQ are not interchangeable.

Our results also provide important implications for managers. First, customers develop overall judgments about e-service providers by evaluating some critical quality characteristics. Therefore, firms must regularly monitor their website performance along these dimensions and take corrective actions as needed. Furthermore, customer feedback analysis provided in the form of comments, service evaluations, or complaints on the focal website, third-party sites, and online forums can serve as valuable sources of information. The use of competitor benchmarking can also help e-service providers to assess their performance on each aspect of service delivery. Our findings highlight that informativeness contributes the most to the overall quality perceptions, followed by fulfillment/reliability and website design. Although online firms must devote special effort and resources to improving these aspects, they should remain mindful of customers' unwillingness to compromise acceptable quality levels. For instance, customers expect an e-service provider to protect them from financial fraud and safeguard their personal information. Even though superior security/privacy features may not appear as offering a competitive advantage, it should be considered "par for the course" for all online service providers and its absence should be considered a competitive disadvantage. In sum, given the customers' different priorities and varying expectation regarding quality dimensions, each online firm must adopt and use its own performance evaluation approach, while developing an in-depth understanding of the perceptions and behavior of different customer segments.

Second, our findings establish that e-SQ has a strong positive effect on trusting beliefs which, in turn, significantly influences customer loyalty. Maintaining a large number of loyal customers is the ultimate objective of all firms that desire to achieve superior performance through online activities. Given the dynamic nature of the Web, meeting customers' expectations is in a constant state of flux. Services-related firms need to invest the necessary resources to continuously monitor the online customers' perceptions and their behavior, anticipate

competitive moves, and monitor technological developments. Timely response to customers' changing needs and preferences is a critical factor for success in online environments. Our findings clearly demonstrate that e-service providers have a great opportunity to acquire and maintain loyal customers by offering a superior online experience.

Limitations and Directions for Future Research

Our findings are influenced by a number of contextual factors which may limit their use in other settings. First, the study was conducted in the context of e-service providers in the travel industry. Although the proposed model of the antecedents and consequences of e-SQ can be readily adapted for use in other contexts, its external validity should be reassessed through replication studies. Second, our focus was on hotel reservation services and, as such, we did not consider hedonic quality aspects or dimensions that relate to the delivery of physical goods, which are important to the online purveyors of such products. Future studies should expand the conceptual framework developed in this study to include such quality elements as well. Third, our data collection was based on a cross-sectional survey, which inhibits making causal inferences between e-SQ and its antecedents and outcomes. For a better understanding of the proximal antecedents and performance consequences of e-SQ, longitudinal studies using causal research designs are needed. Fourth, although on average our response rates compare favorably with those obtained in other online surveys, our samples were relatively small and future studies should use larger samples for enhanced statistical power. Fifth, our informants made a hotel reservation during the year preceding each study and we did not identify the exact timing of their last reservation. Moreover, although we explicitly instructed respondents to complete the questionnaire focusing on their last online booking experience, customers' memory can be circumscribed and thus less than accurate information regarding the site and/or experience may

have been conveyed. Finally, our samples included only people that made a hotel reservation using a specific website. However, it would also be useful to examine the attitudes and perceptions of people who visited a specific website but decided not to conduct any transactions, for various reasons. Unfortunately, contact details of website visitors that are not turned into buying customers are not generally available to online retailers or service providers. More sophisticated tracking techniques may be used for this purpose.

References

Aladwani, Adel M., and Prashant C. Palvia (2002), "Developing and validating an instrument for measuring user-perceived web quality," *Information and Management,* 39, 6, 467-476.

Anderson, Rolph E., and Srini S. Srinivasan (2003), "E-satisfaction and e-loyalty: a contingency framework," *Psychology and Marketing, 20,* 2, 123-138.

Anderson, James C., and David W. Gerbing (1988), "Structural Equation Modeling in Practice: A Review and Recommended Two-step Approach," *Psychological Bulletin, 103,* 3, 411.

Barnes, Stuart J., and Richard T. Vidgen (2002), "An Integrative Approach to the Assessment of e-commerce Quality," *Journal of Electronic Commerce Research,* 3, 3, 114-127.

Barrutia, Jose M., and Ainhize Gilsanz (2013), "Electronic Service Quality and Value Do Consumer Knowledge-related Resources Matter?" *Journal of Service Research, 16,* 2, 231-246.

Bart, Yakov, Venkatesh Shankar, Fareena Sultan, and Glen L. Urban (2005), "Are the Drivers and Role of Online Trust the Same for All Web Sites and Consumers? A Large-scale Exploratory Empirical Study," *Journal of Marketing, 69,* 4, 133-152.

Bauer, Hans H., Tomas Falk, and Maik Hammerschmidt (2006), "eTransQual: a Transaction Process-Based Approach for Capturing Service Quality in Online Shopping," *Journal of Business Research, 59,* 7, 866-875.

Bentler, Peter M. (1995), *EQS structural equations program manual.* Encino, CA: Multivariate Software, Inc.

Brady, Michael K., and J. Joseph Cronin Jr. (2001), "Some New Thoughts on Conceptualizing Perceived Service Quality: A Hierarchical Approach," *Journal of Marketing, 65,* 3, 34-49.

Cai, Shaohan, and Minjoon Jun (2003), "Internet Users' Perceptions of Online Service Quality: A Comparison of Online Buyers and Information Searchers," *Managing Service Quality: An International Journal, 13,* 6, 504-519.

Carlson, Jamie, and Aron O'Cass (2010), "Exploring the Relationships between E-Service Quality, Satisfaction, Attitudes and Behaviours in Content-Driven E-Service Web Sites," *Journal of Services Marketing, 24,* 2, 112-127.

Chang, Hsin Hsin, and Su Wen Chen (2008), "The Impact of Customer Interface Quality, Satisfaction and Switching Costs on E-Loyalty: Internet Experience as a Moderator," *Computers in Human Behavior, 24, 6,* 2927-44.

Chen, Jun, and Sally Dibb (2010), "Consumer Trust in the Online Retail Context: Exploring the Antecedents and Consequences," *Psychology and Marketing, 27,* 4, 323-346.

Collier, Joel E., and Carol C. Bienstock (2006), "Measuring Service Quality in E-Retailing," *Journal of Service Research, 8,* 3, 260-275.

Cristobal, Eduard, Carlos Flavian, and Miguel Guinaliu (2007), "Perceived e-service Quality (Pesq) Measurement Validation and Effects on Consumer Satisfaction and Web Site Loyalty," *Managing Service Quality: An International Journal,* 17, 3, 317-340.

Diamantopoulos, Adamantios, and Heidi M. Winklhofer (2001), "Index Construction with Formative Indicators: An Alternative to Scale Development," *Journal of Marketing Research, 38,* 2, 269-277.

Doney, Patricia M., and Joseph P. Cannon (1997), "Trust in Buyer-Seller Relationships," *Journal of Marketing, 61, 2,* 35-51.

Edwards, Jeffrey R. (2011), "The Fallacy of Formative Measurement," *Organizational Research Methods, 14,* 2, 370-388.

Fan, Xitao, Bruce Thompson, and Lin Wang (1999), "Effects of Sample Size, Estimation Methods, and Model Specification on Structural Equation Modeling Fit Indexes," *Structural Equation Modeling: A Multidisciplinary Journal, 6,* 1, 56-83.

Fassnacht, Martin, and Ibrahim Koese (2006), "Quality of Electronic Services: Conceptualizing and Testing a Hierarchical Model," *Journal of Service Research, 9,* 1, 19-37.

Feldman, Jack M., and John G. Lynch (1988), "Self-generated Validity and Other Effects of Measurement on Belief, Attitude, Intention, and Behavior," *Journal of Applied Psychology, 73,* 3, 421-435.

Finn, Adam, Luming Wang, and Tema Frank (2009), "Attribute Perceptions, Customer Satisfaction and Intention to Recommend E-Services," *Journal of Interactive Marketing, 23,* 3, 209-220.

Fornell, Claes, and David F. Larcker (1981), "Structural Equation Models with Unobservable Variables and Measurement Error: Algebra and Statistics," *Journal of Marketing Research,* 18, 3, 382-388.

Francis, Julie E., and Lesley White (2002), "PIRQUAL: A Scale for Measuring Customer Expectations and Perceptions of Quality in Internet Retailing," *K. Evans and L. Scheer (Eds.),* 263-270.

Gerbing, David W., and James C. Anderson (1988), "An Updated Paradigm for Scale Development Incorporating Unidimensionality and Its Assessment," *Journal of Marketing Research,* 25, 2, 186-192.

Gounaris, Spiros, and Sergios Dimitriadis (2003), "Assessing Service Quality on the Web: Evidence from Business-To-Consumer Portals," *Journal of Services Marketing,* 17, 5, 529-548.

Gupta, Reetika, and Sertan Kabadayi (2010), "The Relationship Between Trusting Beliefs and Web Site Loyalty: The Moderating Role of Consumer Motives and Flow," *Psychology and Marketing, 27,* 2, 166-185.

Ha, Sejin, and Leslie Stoel (2009), "Consumer e-shopping Acceptance: Antecedents in a Technology Acceptance Model," *Journal of Business Research, 62,* 5, 565-571.

Harris, L. C., and Goode, M. H. (2004), "The Four Levels of Loyalty and the Pivotal Role of Trust: A Study of Online Service Dynamics," *Journal of Retailing, 80,* 2, 139-158.

Harris, Lloyd C., and Mark MH Goode (2008), "Satisfiers and Dissatisfiers in the Online Environment: A Critical Incident Assessment," *Journal of Service Research, 10,* 4, 347-364.

Janda, Swinder, Philip J. Trocchia, and Kevin P. Gwinner (2002), "Consumer Perceptions of Internet Retail Service Quality," *International Journal of Service Industry Management, 13,* 5, 412-431.

Jarvis, Cheryl Burke, Scott B. MacKenzie, and Philip M. Podsakoff (2003), "A Critical Review of Construct Indicators and Measurement Model Misspecification in Marketing and Consumer Research," *Journal of Consumer Research, 30,* 2, 199-218.

Kaynama, Shohreh A., and Christine I. Black (2000), "A Proposal to Assess the Service Quality of Online Travel Agencies: An Exploratory Study," *Journal of Professional Services Marketing, 21,* 1, 63-88.

Kim, Jiyoung, Byoungho Jin, and Jane L. Swinney (2009), "The Role of Etail Quality, E-Satisfaction and E-Trust in Online Loyalty Development Process," *Journal of Retailing and Consumer Services, 16,* 4, 239-247.

Kim, Soyoung, and Leslie Stoel (2004), "Apparel Retailers: Website Quality Dimensions and Satisfaction," *Journal of Retailing and Consumer Services, 11,* 2, 109-117.

Ladhari, Riadh (2010), "Developing e-Service Quality Scales: A Literature Review," *Journal of Retailing and Consumer Research, 17,* 6, 464-477.

Law, Kenneth S., and Chi-Sum Wong (1999), "Multidimensional constructs M structural equation analysis: An illustration using the job perception and job satisfaction constructs," *Journal of Management,* 25, 2, 143-160.

Lee, Gwo-Guang, and Hsiu-Fen Lin (2005), "Customer Perceptions of E-Service Quality in Online Shopping," *International Journal of Retail and Distribution Management, 3*3, 2, 161-176.

Lewis, Bruce R., Gary F. Templeton, and Terry Anthony Byrd (2005), "A Methodology for Construct Development in MIS Research," *European Journal of Information Systems, 14, 4,* 388-400.

Li, Y.N., Tan, K.C. and Xie, M. (2002), "Measuring Web-Based Service Quality," *Total Quality Management, 13,* 5, 685-700.

Lin, Jiun-Sheng Chris, and Pei-Ling Hsieh (2011), "Assessing the Self-Service Technology Encounters: Development and Validation of SSTQUAL Scale," *Journal of Retailing, 87,* (2), 194-206.

Long, Mary, and Charles McMellon (2004), "Exploring the Determinants of Retail Service Quality on the Internet," *Journal of Services Marketing, 18,* 1, 78-90.

MacKenzie, Scott B., Philip M. Podsakoff, and Cheryl Burke Jarvis (2005), "The Problem of Measurement Model Misspecification in Behavioral and Organizational Research and Some Recommended Solutions," *Journal of Applied Psychology, 90,* 4, 710-730.

MacKenzie, Scott B., Philip M. Podsakoff, and Nathan P. Podsakoff (2011), "Construct Measurement and Validation Procedures in MIS and Behavioral Research: Integrating New and Existing Techniques," *MIS Quarterly, 35,* 2, 293-334.

Madu, Christian N., and Assumpta A. Madu (2002), "Dimensions of E-Quality," *International Journal of Quality and Reliability Management,* 19, 3, 246-258.

Mayer, Roger C., James H. Davis, and F. David Schoorman (1995), "An Integrative Model of Organizational Trust," *Academy of Management Review, 20,* 3, 709-734.

Montoya-Weiss, Mitzi M., Glenn B. Voss, and Dhruv Grewal (2003), "Determinants of Online Channel Use and Overall Satisfaction with a Relational, Multichannel Service Provider," *Journal of the Academy of Marketing Science, 31,* 4, 448-458.

Moorman, Christine, Rohit Deshpande, and Gerald Zaltman (1993), "Factors Affecting Trust in Market Research Relationships," *Journal of Marketing, 57,* 1, 81-101.

Morgan, Robert M., and Shelby D. Hunt (1994), "The Commitment-Trust Theory of Relationship Marketing," *Journal of Marketing, 58,* 3, 20-38.

Oliver, Richard L. (1999), "Whence Consumer Loyalty?" *Journal of Marketing, 63,* 1, 33-44.

Parasuraman, Ananthanarayanan, Valarie A. Zeithaml, and Arvind Malhotra (2005), "E-S-QUAL: A Multiple-Item Scale for Assessing Electronic Service Quality," *Journal of Service Research, 7,* 3, 213-233.

Rossiter, John R. (2002), "The C-OAR-SE Procedure for Scale Development in Marketing," *International Journal of Research in Marketing, 19,* 4, 305-335.

Rossiter, John R. (2007), "Toward a Valid Measure of E-Retailing Service Quality," *Faculty of Commerce-Papers, 2,* 3, 36-48.

Rossiter, John R. (2009), "ER-SERVCOMPSQUAL: A Measure of E-Retailing Service Components Quality," *Service Science, 1,* 4, 212-224.

Rowley, Jennifer (2006), "An Analysis of the E-Service Literature: Towards a Research Agenda," *Internet Research, 16,* 3, 339-359.

Schlosser, Ann E., Tiffany Barnett White, and Susan M. Lloyd (2006), "Converting Web Site Visitors into Buyers: How Web Site Investment Increases Consumer Trusting Beliefs and Online Purchase Intentions," *Journal of Marketing, 70,* 2, 133-148.

Shankar, Venkatesh, Amy K. Smith, and Arvind Rangaswamy (2003), "Customer Satisfaction and Loyalty in Online and Offline Environments," *International Journal of Research in Marketing, 20,* 2, 153-175.

Sharma, Subhash, Srinivas Durvasula, and William R. Dillon (1989), "Some Results on the Behavior of Alternate Covariance Structure Estimation Procedures in the Presence of Non-Normal Data," *Journal of Marketing Research (JMR), 26,* 2, 214-221.

Sohn, Changsoo, and Suresh K. Tadisina (2008), "Development of e-Service Quality Measure for Internet-Based Financial Institutions," *Total Quality Management,* 19, 9, 903-918.

Srinivasan, Srini S., Rolph Anderson, and Kishore Ponnavolu (2002), "Customer Loyalty in E-Commerce: An Exploration of its Antecedents and Consequences," *Journal of Retailing*, 78, 1, 41-50.

Tan, Chee-Wee, Izak Benbasat, and Ronald T. Cenfetelli (2013), "IT-mediated Customer Service Content and Delivery in Electronic Governments: An Empirical Investigation of the Antecedents of Service Quality," *MIS Quarterly, 37,* 1, 77-109.

Wilcox, James B., Roy D. Howell, and Einar Breivik (2008), "Questions About Formative Measurement," *Journal of Business Research, 61,* 12, 1219-1228.

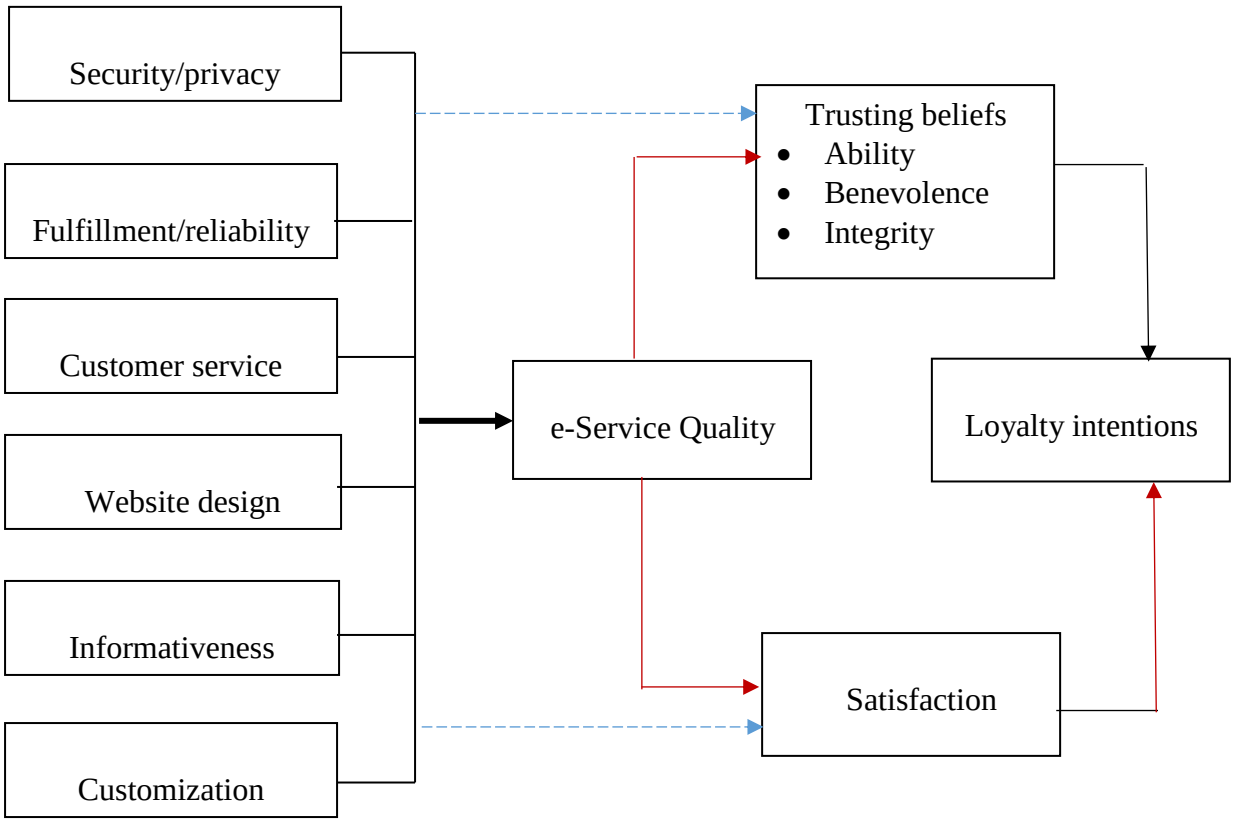
Wolfinbarger, Mary, and Mary C. Gilly (2003), "eTailQ: Dimensionalizing, Measuring and Predicting Etail Quality," *Journal of Retailing, 79,* 3, 183-198.

Yang, Zhilin, Minjoon Jun, and Robin T. Peterson (2004), "Measuring Customer Perceived Online Service Quality: Scale Development and Managerial Implications," *International Journal of Operations and Production Management, 24,* 11, 1149-1174.

Yoo, Boonghee, and Naveen Donthu (2001), "Developing a Scale to Measure the Perceived Quality of an Internet Shopping Site (SITEQUAL)," *Quarterly Journal of Electronic Commerce, 2,* 1, 31-45.

Zeithaml, Valarie A., Arun Parasuraman, and Arvind Malhotra (2002), "Service Quality Delivery through Web Sites: A Critical Review of Extant Knowledge," *Journal of the Academy of Marketing Science, 30,* 4, 362-375.

Figure 1
A conceptual model of the proximal antecedents and customer outcomes of e-service quality



Dashed lines represent the additional relationships estimated for the partial-mediation model

Table 1
Measurement Scales, Confirmatory Factor Analysis Results, and Reliabilities

Constructs and Measurement Items	Standardized Loadings[a] Study 2	Study 3
Security/Privacy[c] (Wolfinbarger and Gilly 2003)		
(α = .92/.89; CR = .92/.89; AVE = .70/.73)		
I feel safe in my transactions with this website	.74[b]	---
I feel like my privacy is protected at this site	.84 (12.69)	.85[b]
I trust this site will not misuse my personal information	.82 (12.29)	.89 (15.00)
I feel I can trust this website	.91 (13.72)	.82 (13.40)
The website instils confidence in customers	.85 (12.88)	---
Fulfillment/Reliability[c] (Wolfinbarger and Gilly 2003)		
(α = .90/.91; CR = .90/.92; AVE = .70/.74)		
You get what you booked from this site	.80[b]	.82[b]
This website gets bookings correct	.87 (14.47)	.91 (15.60)
The online receipt informs me of the total charges that will be debited against my credit card	.83 (13.55)	.79 (12.70)
Transactions at this website are error-free	.86 (14.25)	.91 (15.72)
Customer Service[c] (Wolfinbarger and Gilly 2003)		
(α = .87/.84; CR = .87/.86; AVE = .68/.67)		
The website is ready and willing to respond to customer needs	.80[b]	.63[b]
Customer service personnel are always willing to help you	.82 (13.59)	.89 (9.67)
Inquiries are answered promptly	.86 (14.39)	.92 (9.81)
Informativeness[c] (Wolfinbarger and Gilly 2003)		
(α = .91/.93; CR = .91/.93; AVE = .78/.82)		
At this site, I have the full information at hand	.88[b]	.93[b]
The website provides in-depth information	.91 (18.74)	.94 (23.32)
The site helps me research services	.87 (17.21)	.85 (17.93)
Website Design[c] (Bart et al. 2005; Montoya-Weiss, Voss, and Grewal 2003)		
(α = .94/.88; CR = .94/.88; AVE = .71/.72)		
The site offers a logical layout that is easy to follow	.76[b]	---
The illustrations for the services at the site are helpful in making a purchase decision	.83 (13.19)	---
The site provides a clear directory of available services	.78 (12.29)	---
The site is visually appealing	.91 (14.67)	.80[b]
The visual appearance and manner of the site is professional	.90 (14.53)	----
The site displays a high level of artistic sophistication/creativity	.86 (13.74)	.86 (13.05)
The site is engaging and captures attention	.83 (13.25)	.88 (13.29)
Customization[c] (Wolfinbarger and Gilly 2003; Srinivasan, Anderson, and Ponnavolu 2002)		
(α = .96/.93; CR = .96/.93; AVE = .75/.69)		
This website enables me to choose services that are tailor-made for me	.82[b]	---
The advertisements and promotions that this website sends to me are tailored to my situation	.79 (13.85)	---
This website makes me feel that I am a unique customer	.87 (16.12)	.89[b]
This website gives me personal attention	.90 (16.93)	.91 (18.90)
This website understands my specific needs	.93 (17.92)	.83 (15.74)
This site has features that are personalized for me	.90 (16.77)	.87 (16.91)
This website stores all my preferences and offers me extra services or information based on my preferences	.85 (15.31)	.72 (12.14)
This site does a pretty good job guessing what kind of things I might want and making suggestions	.86 (15.66)	.74 (12.82)
e-Service Quality[c] (Montoya-Weiss, Voss and Grewal 2003)		
(α = .93/.92; CR =.93/.92; AVE = .78/.79)		
This site provides a high level of overall service	.90[b]	.91[b]
This site provides convenient service	.92 (21.46)	.86 (17.12)
This site provides reliable service	.89 (20.11)	.91 (19.54)
This site provides helpful assistance	.83 (16.92)	---

Table 1 continued

Constructs and Measurement Items	Standardized Loadings [a] Study 3
Trusting beliefs (Schlosser, White, and Lloyd 2006)	
A. Ability[c] (α = .92; CR = .92; AVE = .67)	**.62 (7.26)**
This website seems very capable of performing online transactions	.80[b]
This website appears to be successful at the things it tries to do	.74 (11.14)
This website seems to have much knowledge about what needs to be done to fulfill online transactions	.88 (13.90)
I feel very confident about this website's online skills	.87 (13.87)
This website appears to have specialized capabilities that can increase its performance with online transactions	.78 (11.97)
This website appears to be well qualified in the area of e-commerce	.83 (12.93)
B. Benevolence[c] (α = .93; CR = .93; AVE = .81)	**.80 (11.04)**
This website seems very concerned about my welfare	.88[b]
My needs and desires appear to be important to this website	.96 (20.28)
This website seems to really look out for what is important to me	.86 (16.70)
C. Integrity[c] (α = .90; CR = .90; AVE = .70)	**.89** [b]
This website seems to have a strong sense of justice	.85[b]
This website appears to try hard to be fair in dealings with others	.84 (14.08)
I like this website's values	.85 (14.45)
Sound principles seem to guide this website's behavior	.81 (13.44)
Satisfaction[c] (Wolfinbarger and Gilly 2003) (α = .85; CR = .86; AVE = .67)	
I am sure it was the most right thing to make my travel arrangements at this website	.74 [b]
I am satisfied with my decision to make my travel arrangements at this website	.82 (11.55)
I am happy I made my travel arrangements at this website	.88 (12.49)
Loyalty Intentions[c] (Parasuraman, Zeithaml, and Malhotra 2005) (α = .92; CR = .93; AVE = .72)	
I say positive things about this website to other people	.93[b]
I will recommend this website to someone who will ask my advice	.96 (25.71)
I encourage friends and others to do business with this website	.88 (19.89)
I consider this website to be my first choice for future transactions	.79 (15.44)
I will do more business with this website in the coming months	.65 (10.94)

[a] *t*-values are in parentheses.
[b] Item fixed to set the scale..
[c] Seven-point scale, anchored by "Strongly Disagree" and "Strongly Agree."

Table 2
Descriptive statistics and correlation matrix[1]

	1	2	3	4	5	6	7	8	9	10	11	12
1. Security/privacy	---											
2. Fulfillment/reliability	.62	---										
3. Customer service	.52	.57	---									
4. Informativeness	.60	.69	.68	---								
5. Website design	.50	.42	.55	.59	---							
6. Customization	.44	.22	.49	.48	.61	---						
7. e-service quality	.57	.62	.58	.72	.56	.46	---					
8. Ability	.39	.47	.43	.53	.39	.29	.64	---				
9. Benevolence	.33	.16	.42	.44	.49	.67	.35	.35	---			
10. Integrity	.39	.29	.50	.49	.48	.58	.45	.51	.70	---		
11. Satisfaction	.51	.55	.54	.66	.52	.45	.70	.65	.48	.57	---	
12. Loyalty intentions	.41	.43	.47	.59	.45	.48	.57	.61	.51	.62	.71	---
Mean Score	5.35	6.22	5.48	5.73	5.04	4.24	5.80	5.92	4.46	4.88	5.59	5.48
Standard Deviation	1.17	1.02	1.12	1.15	1.14	1.37	1.06	.94	1.39	1.08	1.10	1.14

[1]Correlations and descriptive statistics are based on the mean scores of the items comprising each construct.
[2]All correlations are significant at $p<.05$.

Table 3
Structural Model Estimation Results for the Proximal Antecedents of e-SQ
Full-mediation model[1]

Hypothesized Paths	Expected sign	Standardized coefficient	*t*-value	*p*-value
H1a Security/Privacy→ e-Service Quality	+	.13	2.34	.020
H1b Fulfillment/Reliability → e-Service Quality	+	.37	6.27	.000
H1c Customer service → e-Service Quality	+	.04	.71	.478
H1d Informativeness → e-Service Quality	+	.71	11.18	.000
H1e Website design → e-Service Quality	+	.11	1.93	.054
H1f Customization → e-Service Quality	+	.15	2.82	.005
H2 e-Service Quality → Trusting beliefs	+	.52	4.97	.000
H3 e-Service Quality → Satisfaction	+	.65	7.18	.000
H4 Trusting beliefs → Loyalty intentions	+	.27	4.09	.000
H5 Satisfaction → Loyalty intentions	+	.74	8.79	.000

[1]Fit statistics for structural model $\chi^2_{(961)} = 2013.91$, $p<.00$; $\chi^2/df = 2.10$; CFI=.97; NNFI=.97; RMSEA = 0.069.

Table 4
Structural Model Estimation Results when e-Service Quality is Specified as a Second-order Reflective Scale[1]

Hypothesized Paths	Expected Sign	Standardized Coefficient	*t*-value	*p*-value
H2 e-Service Quality → Trusting beliefs	+	.76	5.98	.000
H3 e-Service Quality → Satisfaction	+	.76	7.30	.000
H4 Trusting beliefs → Loyalty intentions	+	.22	3.46	.000
H5 Satisfaction → Loyalty intentions	+	.77	9.69	.000

[1]Fit statistics for structural model $\chi^2_{(847)} = 1923.70$; $p < .00$; $\chi^2/df = 2.27$; CFI = .97; NNFI = .96; RMSEA = 0.075.

APPENDIX 1
An overview of the most representative studies of Electronic Service (e-Service) Quality Conceptualization and Operationalization

Author(s)	Domain	No of Dimensions**	Dimensions***						Outcome variables
			Security/ Privacy *(64%)*	Fulfilment/ Reliability *(64%)*	Customer Service *(72%)*	Website Design *(72%)*	Informativeness *(44%)*	Customization *(24%)*	
Aladwani and Palvia (2002)	Multi-context	4				x	x		overall quality
Barnes and Vidgen (2002) *(WEBQUAL 4.0)*	e-Bookstores	3			x		x		-
Bauer et al. (2006) *(eTransQual)*	Online retailing	5		x	x	x			perceived value customer satisfaction
Cai and Jun (2003)	Online retailing	4	x	x	x	x			overall quality
Collier and Bienstock (2006) *(E-Service Quality)*	Online retailing	12	x	x	x	x	x		satisfaction behavioral intentions
Cristobal et al. (2007) - *(PeSQ)*	e-Commerce	4		x	x	x			satisfaction website loyalty
Fassnacht and Koese (2006)	e-Service	9		x		x	x		overall quality
Francis and White (2002) *(PIRQUAL)*	Online retailing	6	x		x	x			behavioral intentions
Gounaris and Dimitriadis (2003)	Web portals	3			x		x		-
Holloway and Beatty (2008)	Online retailing	4	x	x	x	x			satisfaction, repurchase intention, WOM
Janda et al. (2002)	Online retailing	5	x				x		satisfaction, WOM, likelihood of future purchases, likelihood of complaining
Kaynama and Black (2000) *(e-Qual)*	Online travel agents	7			x	x		x	-

Kim and Stoel (2004)	Online retailing	6	x			x	x			x	satisfaction
Ladhari (2010)*	e-Service	6	x	x	x	x	x				-
Lee and Lin (2005)	e-Bookstores	5	x	x	x	x				x	overall service quality, satisfaction, purchase intentions
Li et al. (2002)	e-Service	6				x		x			-
Long and McMellon (2004)	Online retailing	5			x	x					overall quality, positive WOM
Madu and Madu (2002)*****	e-Quality	15	x	x	x	x				x	-
Parasuraman et al. (2005) *(E-S-Qual* and *E-RecS-Qual)*	e-Service	7	x	x	x						perceived value, loyalty intentions
Rowley (2006)*****	e-Service	10	x	x		x	x			x	-
Sohn and Tadisina (2008)	Online financial services	6	x	x		x	x			x	-
Wolfinbarger and Gilly (2003) *(eTailQ)*	Online retailing	4	x	x	x	x					overall quality, satisfaction, loyalty intentions, attitude towards the website
Yang et al. (2004)	Online banking	6	x	x	x						overall quality, satisfaction
Yoo and Donthu (2001) *(SITEQUAL)*	Online retailing	4	x			x					overall quality, attitude towards the website, site loyalty, site equity, purchase intention, site revisit intentions
Zeithaml et al. (2002)* *(e-SQ)*	e-Service	5	x	x		x	x				-

*Conceptual papers based on literature reviews – Dimensions appearing most consistently among papers reviewed.

**There were other dimensions that were not as frequently appearing in order to include in this table (i.e., ease of use, accessibility, sensation, enjoyment, ownership conditions, competence, tangibility, empathy, etc.).

***Frequency of appearance of each dimension in the studies in parenthesis.

APPENDIX 2
Items dropped following scale purification procedures

Security/Privacy

- The website has adequate security features
- I feel secure giving out credit card information at this website
- The website clearly explains how user information is used

Fulfillment/Reliability

- The service I received was represented accurately by this website
- Easy booking and payment mechanisms exist

Customer Service

- When you have a problem, the website shows a sincere interest in solving it

Informativeness

- This website has comprehensive information
- This website is a very good source of information

Customization

- This website makes booking recommendations that match my needs
- A first time visitor can make a booking through this website without much help

Website design

- The download time was acceptable
- All texts and menus displayed properly
- Information on the site can be obtained quickly
- The site content is easy for me to understand
- The overall layout of the website is clear
- I like the graphics of this website